Comments of Mike Philbin:
Hydrology and Riparian Program Lead
BLM - Montana/Dakota's State Office

Introduction

It is obvious a lot of hard work has put into this form. These comments are intended to provide recommendations in order to further improve the product from a field and management perspective.

Reviewers Background

Over my 17 year career with the Forest Service and BLM (plus two degrees in aquatic sciences), I've completed over a thousand assessments ranging from rapid PFC(s) to very detailed evaluations utilizing cross sections, macroinvertebrates, and water chemistry. I have also spent several years as an "end user" of inventory data. As such, I have extensive experience in this subject area.

General Thoughts

An evaluation tool needs to address several factors in order to be of value to resource managers. It needs to: (1) provide the flexibility to deal with the wide range of natural variation (both within a wetland and between different wetlands), (2) it needs to emphasize function rather than conditions or causes, (3) it needs to address the key indicators of ecosystem health, (4) it needs to deal with "bottlenecks" as well as cumulative impacts (total scores), (5) it needs to focus on things that can be directly observed, (6) it needs to be kept short, concise and to the point, and (7) it needs to be firmly based in science (defensible). Below I will discuss how the form can be tweaked to better address these points.

Point 1: To deal with variation, a tool should base its rating on departures from either "natural" or a site's own "potential". A rating system that assigns points based on current conditions and then assumes that all systems could achieve the same "maximum rating" will almost certainly mis-categorize some wetlands. This is true of all "index based methods" that do not consider "natural variation". While it seems obvious that a score of .7 for one wetland would not equal a .7 for another; it's inevitable that someone will incorrectly draw this conclusion. The wide range of variation also necessitates that the form be general enough to be widely applicable. As such, it should contain the core elements that can be used in all wetland types and geo-climatic settings. Such a tool is easier to develop when the "bar" upon which the wetland is measured is the sites own potential. *RECOMMENDATION: Rather than assign ratings of excellent, good, fair, and poor based on points; leave this determination to the professional analyst. For example a wetland with a score or .68 may in fact have greater integrity than one rated as .71 (based on which fields received which scores, wetland type, etc…). However, given the draft system, one rated .71 would receive the "better" rating. While the analysts could change the "call", this could lead to problems and even legal challenges. I've seen numerous problems when an approved method says one thing and the interpreter says another. At a minimum, a disclaimer should be added allowing the analyst to override a rating. Another possibility (my preferred alternative) would be to develop a two column form with one column being potential and the other actual. The ratio between the two would produce the rating.*

Point 2: It's more desirable to address function because while an impact may create a short term reduction in condition, a properly functioning system can generally process the input while continuing to support its many uses. Natural fluctuations in conditions are to be expected in all systems and do not necessarily imply serious problems. "Causes," like a dike, may or may not result in adverse effects. By focusing on "causes", we could end up ignoring the natural resiliencies of certain systems. For example, a very resilient system that can handle a particular disturbance may still be "marked down" for a "cause" that is not really affecting function. Mixing "cause and effect" can also produce skewed results as these are often very strongly correlated. *RECOMMENDATION: There are several fields that address causes or conditions as opposed to function in this form. These should be modified keeping the emphasis on actual alterations to function or wetland integrity.*

Point 3: There are many variables that could be assessed in all systems. Rapid assessments should provide basic information that will allow end users to determine the ability of a particular area to support its various uses. They are also useful in prioritizing/scheduling more detailed survey work. If we're looking at the physical, chemical, and biological integrity of a wetland (CWA goals); we need to ask what are the important "universal" elements of each goal? A local or site specific evaluation can get into the specifics later. We also need to be careful not to try to be "everything to everybody." Often researchers have their "pet" areas that they want included. However, to keep a procedure useful and reliable; the examiner/crew must be properly trained in all aspects of the method. If a form tries to address everyone's questions, the experience of the examiner (and therefore the data itself) comes into question. *RECOMMENDATION: There is a lot in this form. I would argue that a rapid assessment is most appropriate for addressing physical and some basic biological elements (vegetation). In many cases, the comment section is the best place to address variables such chemistry, point sources, or detailed biology. If a chemical alterations or special biological conditions are noted, a more detailed assessment would be warranted. In most cases, an analyst can look at the physical and vegetative information and determine whether a more detailed analysis is needed (is the site suitable?).*

Point 4: In addition to the potential for mis-categorizing wetlands, these types of scoring systems can often miss critical deficiencies or problem areas. These "bottlenecks" need to be accounted for. With so many questions and the way the main sections are weighted, key "bottlenecks" can get lost or "masked" by other, often less important variables. This can result in serious problems for both physical and biological systems. An example is a culvert that is impairing sediment transport and acting as a barrier to aquatic organism passage. The current system does not allow for a "fatal flaw" or "bottleneck" to override the rating. It merely averages everything into a "cumulative rating." *RECOMMENDATION: I recommend a process where one really bad variable (in a key area) can toss the entire system into a "problem" category. This can be incorporated by something as simple as a critical "fill in the blank" field where the presence of a "bottleneck" can override the rating. This type of system would allow the user to see what the rating would be if the "bottleneck" were corrected. This could help in prioritizing restoration.*

Point 5: This one's pretty straight forward. We can train the crews in observations, but it takes a lot of education, training, or experience to make a "reasonably valid" interpretation when dealing with these highly complex systems. *RECOMMENDATION: The fields should be limited to*

direct observations. All interpretation should be left to the professionals. Two specific improvements would be to (1) change "evidence of contamination" (WQ Section) to evidence of oil, excessive foaming…" and (2) change impairment (under Hydrogeomorphic - Recovery Trends) to disturbances. Evidence of contamination and impairment carries a much more regulatory meaning.

Point 6: Ideally a rapid assessment would take up the front and back of a page. This particular form is 10 pages long. A form this long is: (1) difficult to carry when working on multiple plots, (2) difficult to use in the field (it's difficult to thumb through ten pages of forms with the wind blowing or when it's raining or it's cold - especially with gloves on), and (3) confusing as pages get mixed up - both in the field and in the office. I believe a rapid assessment should be as short, concise, and simple as possible; while still producing useful and reliable data. A ten page "rapid" assessment would likely result in people rushing through fields without giving enough thought to the questions. This would be more likely in bad weather, when they just want to put their gloves back on! When things are difficult, or people get uncomfortable or frustrated, they can rush reducing the quality of their data. All of this can lead to low consistency between crews. *RECOMMENDATION: The form could be shortened without losing any value. Several questions get at the same thing and others really are not tied to function or even conditions. Ideally the form would use the least amount of fields possible to get a consistent rating (between multiple crews) for each category/major heading. Often this will be in the neighborhood of three to five fields per major category. Changes in format could also help (don't include both species list and fill in the blanks – just have them circle the item in the list.). Good examples of short rapid assessments are the PFC and Pfankuch forms.*

Point 7: The user's guide and documentation are at least as important as the form itself. It will improve consistency and provide the science behind the form. This is critical when facing legal challenges. *RECOMMENDATION: A user's guide should be produced that clearly identifies the rational for all fields. I recommend addressing things like:*

> *Why did you choose a specific variable?*
> *What evidence do you have that it's the best indicator for a CWA or program goal?*
> *What's the science behind the various fields, categories, scores?*
> *How did you come up with the points (the score for a "box")?*
> *How did you establish category breaks (the description in a "box")?*
> *Do all fields with a similar score pose a similar threat to wetland function?*
> *Do you have research that shows that a variable rated 5 is twice as bad as one rated 10?*
> *In determining the final rating, how were the weights established? For example, why is water quality given the least weight?*
> *In all cases include your sources?*

As a management agency, everything we do is challenged. We couldn't use this form without a well documented and strong scientific foundation.

Good examples for documentation include:

TR 1737-**15**: **A User Guide to Assessing Proper Functioning Condition and the Supporting Science for Lotic Areas (1998)**

TR 1737-**16**: **A User Guide to Assessing Proper Functioning Condition and the Supporting Science for Lentic Areas (1999, Revised 2003)**

These are found at: http://www.or.blm.gov/nrst/Tech_References/tech_references.htm

Specific Points

Site Characterization
Why have a section for "Amphibian s and Reptiles" and a species list? Why not just include the list and have them circle what they see? This could shorten the form - a real need.

Water Quality
This section rates cause (the entire column on buffer penetrated by stressor) and effect together. I recommend limiting the "stressor information" to the "top three list" found just below the index block (and in the very last section – overwhelming stressors). We should evaluate functions impacted and not "stressors" that might or might not ever affect the physical, chemical, or biological integrity of the wetland. It also gives real data (instrument readings) less weight than observational information. This real data should fit into the rating in some way. Why are the breaks in blocks set where they are? Are the relative differences between blocks based on anything? For example, does minimal algae growth really warrant the same weight as no toxic contaminants? Does no known point sources or saline seeps within ¼ mile warrant the same score (weight) as no toxic contaminant? Does sediment carry the same risk as petroleum or toxic contaminants (and why are they lumped – wouldn't we want to know which one?)? Do all fields with a score of three pose a similar risk to wetland function?

Finally, whatever score is generated will be "masked" in the final stage when WQ is weighed as only 10% of the total score. As such, a wetland with very poor WQ (even "toxic" or poisonous) could still get a high rating. Given all the calls that must be made to get a water quality rating it's surprising there is so little weight to this section. I recommend the amount of effort used to complete a section be proportionate to the sections weight. For a 10% weight, a simple index such as the following may be appropriate ***(I just made this up so don't read too much into It):***

Water quality does not appear to be reduced due to elevated sediment levels, temperature, nutrients, bacteria, or other pollutants.	Slight change from expected natural condition.	Moderate change from expected natural condition.	Pollutant levels appear to be greatly elevated.
10	8	5	2

Identify Pollutant(s) in order of significance:
Identify Sources:
Identify Method: Measured or Observed

Reduce rating one point for each pollutant over one.

These terms can all be defined in the user guide. The definition may be something like: the stream bed is buried in sediment, water color is altered, algae is chocking channel, …

This method could query on score, pollutants, and sources. These are the same things identified in TMDLs or water quality assessments. Again the matrix is simple because the weight is low. Why spend a lot of time on a variable that really doesn't make that much of a difference in the final rating? In addition, a single simple field could improve consistency between crews.

Hydrogeomorphic

Several of these questions are inter-related. To address this, I think it's important to concentrate on a few key variables. I recommend concentrating on things such as entrenchment or aggradation (vertical instability), excessive bank erosion (lateral instability), and flood plain alteration (accessibility, stability, altered vegetation…). These factors are all important as they indicate altered functions. Currently just the presence of an upstream dam is enough of a reason to reduce a rating, even if no effects are detectable. In short, I feel we should rate the wetland based on impacts and then identify the causes in the comment section.

We should not rate both cause and effect together. The current form mixes cause (dams, dredge and fill, water withdrawals), effect (banks eroding and unstable, entrenchment) and both in the same questions (altered by…). It then evaluates them all together. As such many of the fields are related increasing the possibility that the entire section can be skewed by one type of threat. This is made even more likely as this section keys in on an average of the lowest five scores (and we know the related fields will all get "dinged" to some level).

Do the points make sense? Is bare ground really as important as entrenchment, bank erosion, or rip-raped banks? A shorter, simpler matrix would likely improve consistency between crews.

Since "hydrogeomorphic condition" carries 40% of the final rating, we really need to be careful when putting together this section. The potential problem noted above is magnified since the "hydrogeomorphic condition index" caries so much overall weight.

Hydrogeomorphic Recovery Trends

The block "no impacts are evident…" will be near impossible to achieve. If taken too literally this would make 5 the best score possible for a nice wetland in static condition. An impacted wetland that is slowly improving (but still hammered) would actually rate higher. A slightly impacted wetland with a slight/very slow decline would rate a 0 while a hammered yet static wetland (will remain hammered) would rate a 5. Therefore even though the wetland with a 0 would remain in better condition for the next twenty years, it would rate well below the hammered one. Trend is a tough variable to put a score on. I recommend just noting the trend.

Buffer Condition Assessment

This assessment emphasizes threats and stressors. I recommend that if this section remains in the assessment, that its focus change to function. Key factors influencing how buffers ***function*** are

slope, width, and ground cover or roughness. To account for these factors, I recommend a matrix similar to the following **(*I just made this up so don't read too much into it):***

Moderate Level of Disturbance in Buffer

Slope/Width	0-50 feet	50-100 feet	100-300 feet	300+ feet
0-10	6	8	10	10
10-25	4	6	8	10
25-50	2	4	6	8
50+	0	2	4	6

To incorporate ***conditions,*** the values could be adjusted. For example:

Low Level of disturbance within buffer = Ground Cover >90%, No roads, No active rills or gullies (Add two points up to a maximum of 10)

Moderate Level of Disturbance = Ground Cover between 70-90%, Some logging but less than 50% removal, graveled, paved roads, or grassed over roads (keep scores as in table).

High Level of Disturbance = Ground Cover <70%, Ground based logging, other forms of logging removing >50 of the stand, native surface roads, major highways, rills or gullies. (Subtract 2 points down to a ***minimum*** of 0)

If a road or disturbance prevents the outer portion of the buffer from serving its function, account for it under width. If the feature is affecting how the buffer operates, but the function is still present, account for it under disturbance.

<u>In terms of documentation:</u>
There are numerous references that 300 feet is a very protective buffer. 25% slope also appears to be important. This supports the relative ratings in the 300+ column.

Likewise 50 feet appears to be clearly the least protective supporting the 0-50 column.

The (flat) 0-10 row is well defensible as is the very steep (50+). The middle four are relative to the others. All of this can be supported by current science.

The disturbance factors are also well documented in current literature.

Why are buffers addressed both here and under water quality? It should be in one place or another so it doesn't carry too much weight when determining the final rating.

Vegetation Condition:
I ran out of time and did not give this section much of a review. However, it appears it can be greatly simplified / shortened.

Summary Rating:
Where did the weights come from? For example:

Is water quality really only half as important as buffers (which are intended to protect WQ),
Is water quality only ¼ as important as hydrogeomorphic condition?
Is water quality only 1/3 as important as vegetation?
If water quality is worth so little do we really need to collect so much information? It seems the amount of work required to complete a section should be proportional to the value of that section.

Other Thoughts
The three main criteria for determining "jurisdictional" wetlands are hydrology, soils, and vegetation. This form completely ignores soils. I would recommend adding something as simple as depth of organic horizon, depth to free water, presence of gleying, and presence of mottling to deal with soils. I feel this is a glaring omission in a wetland evaluation.